Award Number: W81XWH-11-1-0737

TITLE: Role of Mitochondrial Inheritance on Prostate Cancer Outcome in African American Men

PRINCIPAL INVESTIGATOR: Martha Mims

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, TX 77030

REPORT DATE: December 2015

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
December 2015	Final	30Sep2011 - 29Sep2015

4. TITLE AND SUBTITLE

Role of Mitochondrial Inheritance on Prostate Cancer Outcome in African American Men

5a. CONTRACT NUMBER
CC

5b. GRANT NUMBER
W81XWH-11-1-0737

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Martha Mims

E-Mail: mmims@bcm.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Baylor College of Medicine
One Baylor Plaza T100
Houston, Texas 77030-3498

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT We are examining the hypothesis that mitochondrial inheritance plays a significant role in aggressiveness of prostate cancer in African Americans. In the first year of the project we identified 2,000 noncancerous tissues samples from African American men with prostate cancer and we have extracted DNA from ~ 1,500 of them. We have validated a robust new DNA sequencing technique developed by our collaborator using single amplicon long-range PCR that permits deep coverage (10,000-20,000X on average) of the mitochondrial genome. We have sequenced 652 samples derived from frozen fully using this technology. The additional DNA samples derived from paraffin embedded (FFPE) tissue were more challenging, but have now been sequenced. Mapping of DNA variants in our sequenced genomes to mitochondrial genes is nearly complete and examples of the data we have obtained are presented here. After months of testing, we determined that the ethidium bromide methodology for generating prostate cancer cell line cybrids was not effective and we have instead used a Rhodamine-6-G procedure. PNT1A cybrid cell lines have been generated and we have characterized these cells. PC3 cells have been more difficult, but we continue to work on this cell line. We anticipate the first manuscripts relating to the large amount of data generated in this project will be written and submitted in the spring of 2016. To our knowledge this represents the largest collection of mitochondrial sequence data from African Americans in existence.

15. SUBJECT TERMS **Mitochondrial DNA sequencing, Prostate Cancer, Cybrid Cells, African Americans**

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	Unclassified	18	USAMRMC

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Page

1. Introduction:

African American men are disproportionately affected by prostate cancer with increased lifetime risk, earlier onset of disease and more advanced stage at diagnosis than Caucasians. Prostate cancer is the leading cause of cancer death in African American men, with mortality more than double that observed in Caucasians. There are nearly 20 million African American men in the US, many of whom face significant risk of developing and dying from prostate cancer. Risk for and aggressiveness of prostate cancer in African American men is thought to originate in part from genetic susceptibility. Several nuclear genes and chromosomal regions have been linked to prostate cancer; however, many studies have not included African American men, and no study has linked genetic polymorphisms with clinical outcome. One factor which has not been carefully examined is mitochondrial inheritance which varies significantly between ethnic and racial groups and could explain large differences in disease characteristics. In the present study, we are examining the hypothesis that mitochondrial inheritance plays a significant role in aggressiveness of prostate cancer in African Americans. Further, we predict that adverse clinical outcomes will be reflected in dysregulation of cellular biochemical processes and in alterations in signaling pathways (Akt pathway, apoptosis). The ability to identify mitochondrial variants or haplogroups that contribute to aggressive disease will help to separate patients with indolent disease who can be spared unnecessary intervention from those who need more immediate and aggressive therapy. Identification of cellular pathways involved will help to target treatment strategies for those with cancers predicted to be more aggressive. In Aim 1 we will conduct a study of mitochondrial inheritance in 1,000 African American men with prostate cancer. We will sequence the mitochondrial genome of all 1,000 samples and determine whether particular mitochondrial variants, genes or haplogroups are associated with markers of aggressive disease (age at diagnosis, stage at diagnosis (including bone metastasis), Gleason score, PSA at diagnosis, PSA recurrence and death from disease). Findings will be replicated using an independent set of 1,000 patients from our own tissue resource with linked clinical data. Our strategy will be to sequence the mitochondrial DNA of all 1,000 patients in the discovery phase as well as in the replication group. We will control for population admixture using the Illumina African American Admixture Panel. Using cybrid technology, we will introduce our previously identified mt10398A variant and mitochondrial variants associated with highly aggressive and least aggressive disease identified in our genotyping study into prostate cell lines (derived from normal and cancer cells). Cybrids differing only in their mitochondrial composition will be examined for viability under metabolic stress, cell cycle distribution, production of reactive oxygen species, O_2 consumption, ATP synthesis, respiratory chain activity and capacity to grow in an anchorage independent manner. The effect of mitochondrial variants on nuclear gene expression will be studied using Western blotting and microarrays.

2. Keywords:

Mitochondrial DNA sequencing, Prostate Cancer, Cybrid Cells, African Americans

3. Accomplishments:

In the text to follow, we provide the Aims and original Statement of Work in italics with progress on the project presented in regular text.

<u>Aim 1:</u> To examine the association between mitochondrial DNA variants and clinical outcome in African American men with prostate cancer.

Task 1: <u>Extraction of all DNA for Initial Study and Validation Set</u>

1a. Complete extraction of DNA/ quantitation of DNA for mitochondrial sequencing *Months 1-9*
(Extraction ongoing at present, expect more than 1,000 samples to be prepared before the project is funded. Additional 1,000 to be extracted for validation set IRB protocol already approved)

We have extracted DNA on more than 1500 Tissue samples from African American Men with Prostate cancer. Work was initially slower than we had initially planned because of difficulties with extracting adequate DNA from some of the FFPE (paraffin-embedded) samples. After significant technical difficulties we have now extracted adequate quantities of DNA from more than 1000 FFPE samples and 652 Frozen tissue samples and these samples have been sent for sequencing.

.

Task 2: <u>Mitochondrial Sequencing</u>

2a. Order and test overlapping mitochondrial primers with FFPE DNA to be sure that all primers work with FFPE tissue DNA	*Ongoing – Month 3*
2b. Redesign/test and mitochondrial primers which do not give good PCR results in 2a.	*Months 4-6*
2c. Establish Database for mitochondrial sequences	*Months 1-5*
2d. PCR mitochondrial sequences for primary study and validation set	*Months 6-20*
2e. Sequence mitochondrial PCR products	*Months 6-20*
(Illumina Admixture Genotyping to be supported by Helis funds, but performed simultaneously)	*Months 6-20*

As reported in our initial project proposal, based on published sequences for mitochondrial PCR primers[1, 2] we designed 61 pairs of overlapping primers to amplify the entire 16.6-kb mitochondrial genome. In order to test this technology, we used 15 matched DNA samples derived from 5 patients – frozen tissue, FFPE tissue and whole genome amplified DNA samples derived from FFPE DNA for each patient. Overall 2/3 of the amplicons had reads in both directions on all 15 samples (including WGA DNA). Five of 61 amplicons failed completely and will be redesigned. Thus, we covered 93% of the genome by reads in at least one direction in 90% of the samples, and by reads in both directions in 84% of the mt genome, an excellent outcome for a first pass analysis over a target. Using these results, we identified 165 variants as compared to the published mitochondrial sequence. Within these variants there was 96-100% concordance of calls for each patient across the three sample types. We have continued to optimize the primers for the Sanger sequencing protocols, but as we reported in our initial application, we have also continued to test new technologies for mitochondrial sequencing. As we reported last year, we have obtained excellent results using a new technique developed by one of the collaborators on this project (Dr. Lee-Jun Wong). This technique enriches the entire human mitochondrial genome by a single amplicon long-range PCR followed by massively parallel sequencing.[3] This protocol utilizes less than 100 ng of tissue DNA and makes possible a one- step approach to provide quantitative base calls, exact deletion junction sequences and quantification of deletion heteroplasmy. As we reported last year, this strategy worked well for our frozen tissue DNA and we have now sequenced 652 samples using this technology.

We were not sure that this strategy will be effective in our FFPE DNA samples. Fortuitiously we had a series of Frozen and FFPE samples derived from the same subject which we could use for comparison purposes. In the figures below we show the result from 4 such Frozen/FFPE pairs. To sequence FFPE samples, we used a sequence-specific capture approach followed by Illumina sequencing. We utlized a custom capture reagent (Roche NimbleGen) targeting the hg19 mitochondrial genome (16.6 Kb). Libraries of ~200 bp insert sizes were prepared from genomic DNA (500 ng) using HGSC Illumina WES protocol and co-captured at 24-48 plex using full-length blockers to capture reads at higher on-target rates. Libraries were pooled and sequenced on a single Illumina HiSeq 2500 lane (~30 Gb). We estimated that each sample would yield 312 Mb of sequence, which roughly would amount to 8000X depth across the target regions. This is a one third less average coverage to the 12000X coverage data currently generated for frozen tissue DNA samples using the long range PCR protocol, where 80 samples are sequenced per lane on Illumina HiSeq. As seen in the figures below, these estimates were relatively accurate with regard to coverage for the FFPE samples. We expected that duplicate rates, which are usually elevated in small target region captures, would not be an issue in this case, as mitochondrial DNA copy number is 1-2 log orders greater than that of the genomic DNA. Non-specific capture of Pseudogenes in the nuclear genome that share high sequence similarity with functional mtDNA genes is another valid concern when using the above described target enrichment approach. However, recent studies have shown that it is possible to bioinformatically determine the limit on heteroplasmy detection due to such contamination and largely eliminate this concern.[4]

Figure 1A shows results from four samples sequenced by the long range primer technology. We now have sequenced 652 samples using this technology. As seen in figure 1A, a distinct advantage provided by amplification of the entire mitochondrial genome by long range PR with a single primer pair is the uniform coverage. This figure demonstrates that we routinely had coverage depth of 10,000-20,000-fold for the mitochondrial genome. This sequencing strategy has multiple advantages including ease of excluding nuclear DNA sequences which are nearly identical to mtDNA and more uniform coverage of the mitochondrial genome because there is much less risk of having a rare or novel variant at the primer binding site when only one set of primers is used. Deep coverage allows much more accurate base calls and improves the limit of detection for heteroplasmy from about 15% with Sanger sequencing to ~ 1-2% with the long range primer technique. Figure 1B shows sequence data from FFPE DNA samples derived from the same 4 patients; these sequences have been performed in duplicate (data not show) with similar results. Note that as we predicted, coverage depth for the FFPE DNA samples was more like 8000X across the target region. Some question has arisen as to why the depth of coverage is not uniform over the entire mitochondrial genome (sawtooth pattern). We feel this may relate to the GC content of the mitochondrial genome which demonstrates a very similar sawtooth pattern, but we do not feel it will affect our overall results.

Figure 1: Mitochondrial DNA sequence data

Figure 1A: Panels A-D show representative coverage for 4 different DNA samples across the entire 16,569 bp mitochondrial genome. The y axis represents fold coverage at each position (x coordinate) along the mitochondrial genome

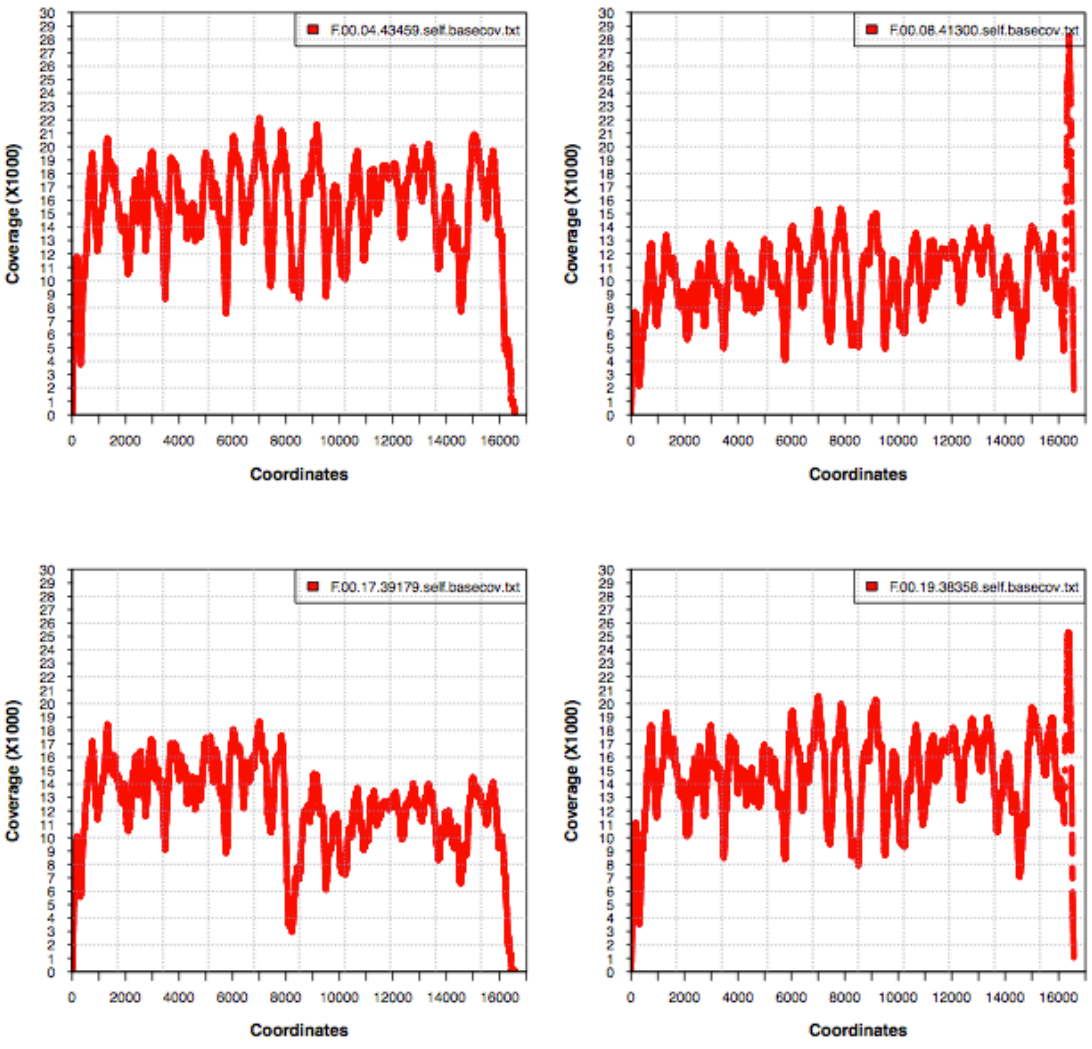
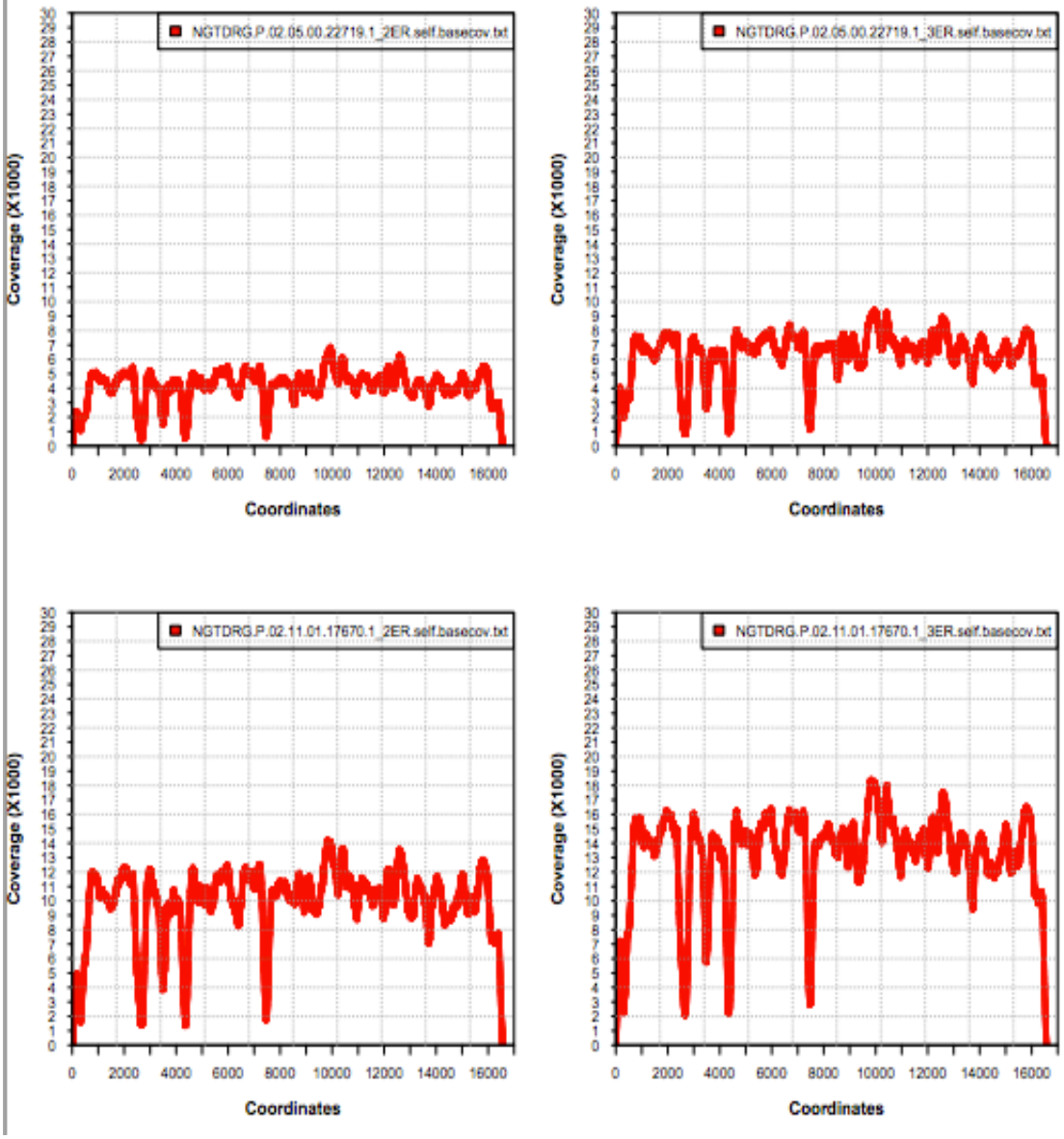


Figure 1B: Panels A-D show representative coverage for 4 different FFPE DNA samples across the entire 16,569 bp mitochondrial genome. Note that these 4 samples are matched to the Frozen DNA samples in panel 1A. The y axis represents fold coverage at each position (x coordinate) along the mitochondrial genome.



Using this data we were able to compare SNP counts in Frozen vs FFPE DNA samples. These data (figure 2) demonstrate that similar numbers of SNPs were called in Frozen and FFPE samples. Figure 3 shows that these SNPs were not completely concordant and we are working on this issue.

Figure 2 Mitochondrial SNP counts in Frozen and FFPE DNA samples derived from the same patient

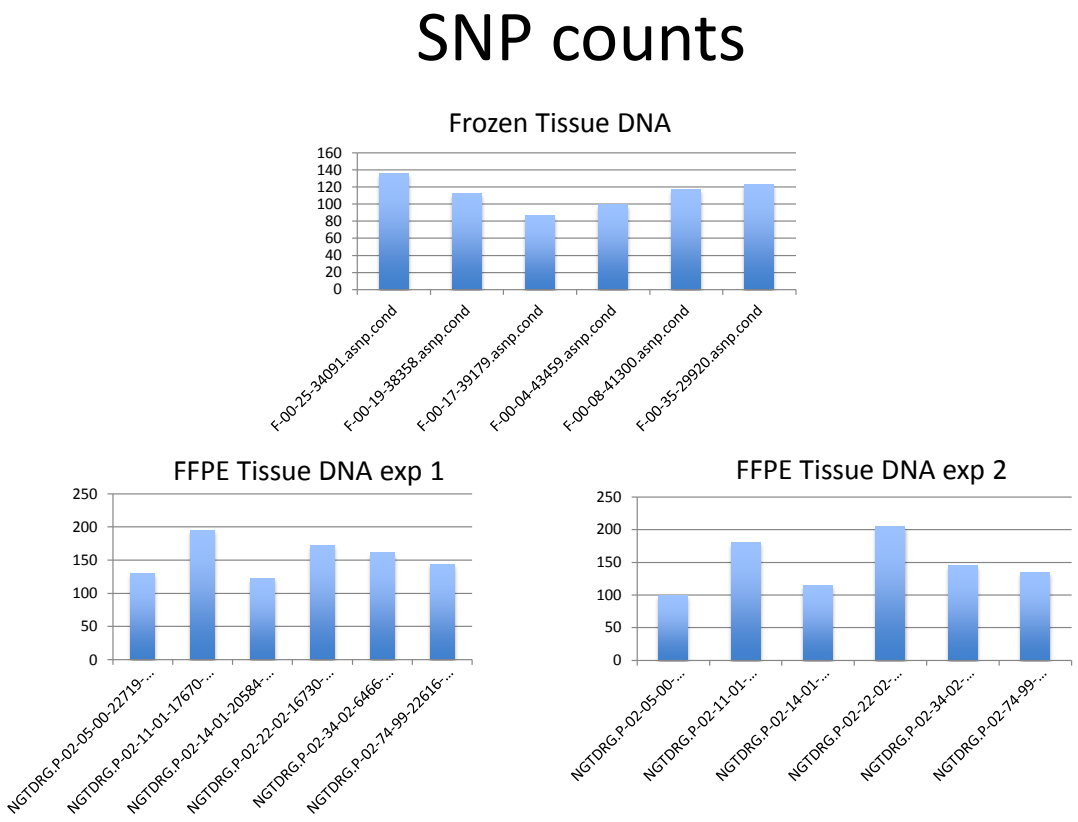
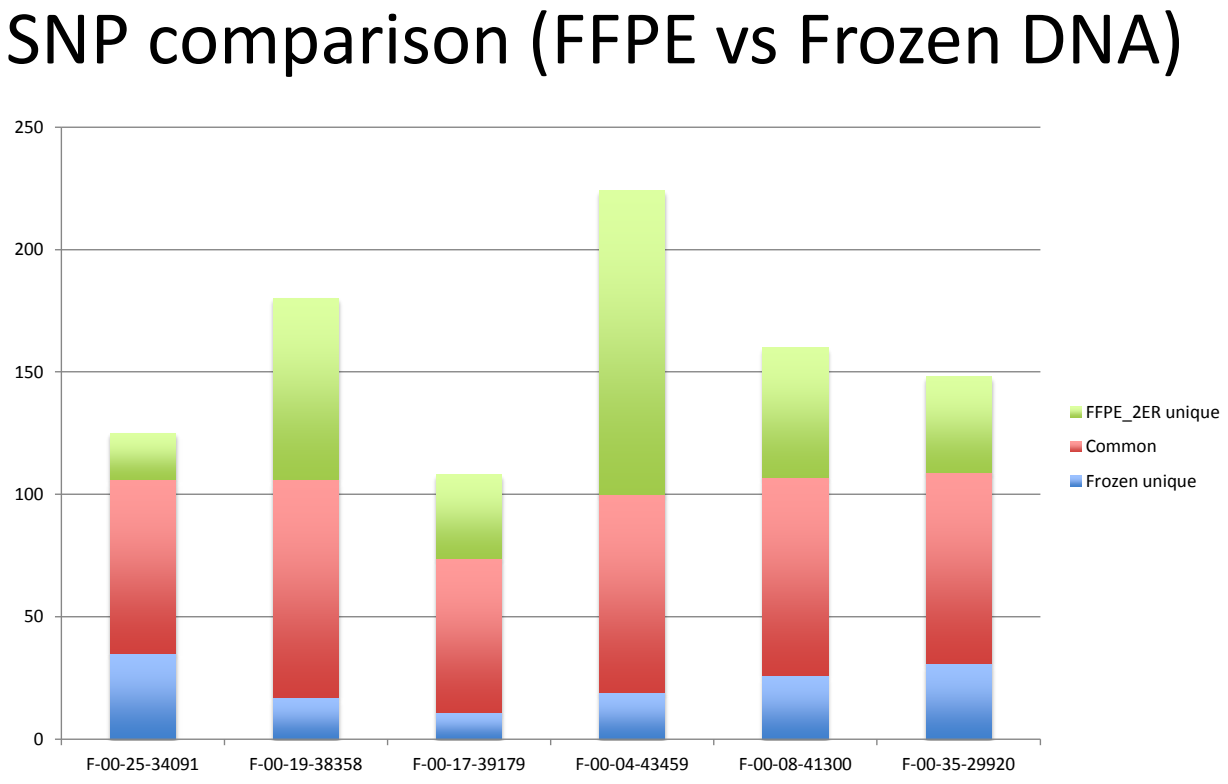


Figure 3 Mitochondrial SNP comparison in Frozen and FFPE DNA samples derived from the same patient



Task 3: Analysis of Mitochondrial Data

3a. Analyze mitochondrial sequences and integrate with clinical data	*Months 20-28*
3b. Admixture analysis based on Illumina data	*Months 20-28*
3c. Identify mitochondrial variants of interest for cybrid study and match with lymphoblast bank	*Months 26-28*

Figure 2A demonstrates the single nucleotide variant (SNV) calls for the first approximately two hundred subjects. On average each subject has about 150 SNVs with the Cambridge mitochondrial sequence used as reference.[5,6] As shown in our previous report, most of the data at each coordinate clustered at 0 or 1. Since the mitochondrial genome is essentially haploid this suggests there there is little contamination with nuclear DNA sequences. Similar data for insertions and deletions is shown in figure 2B.

Figure 2: Single nucleotide variants in the mitochondrial genome for the first sequenced subjects.

Figure 2A: Single nucleotide variants in the mitochondrial genome for the first sequenced subjects.

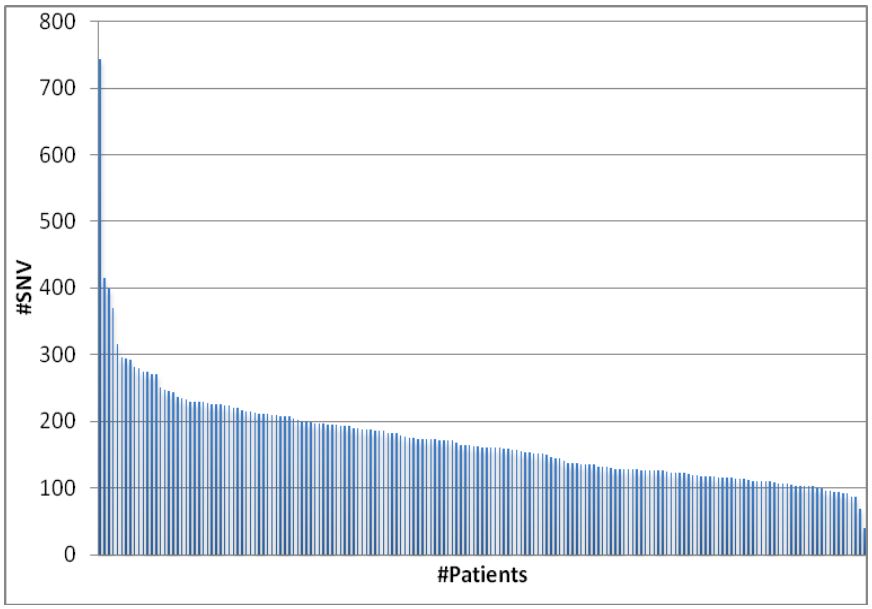
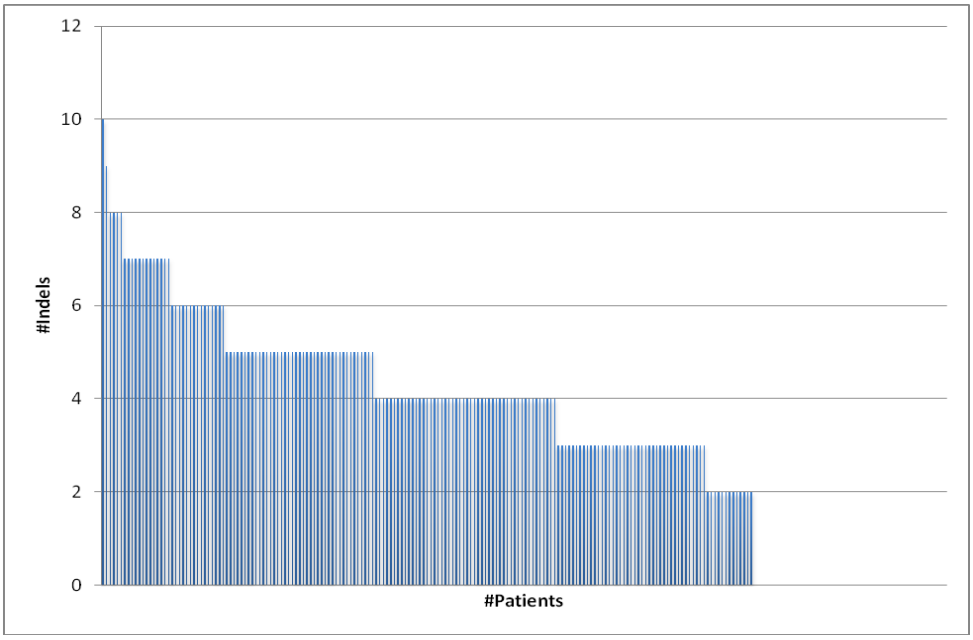


Figure 2B: Insertions and Deletions in the mitochondrial genome for the first sequenced subjects.



We have now created a database of the sequence data which will ultimately include not just mtDNA sequence data, but nuclear SNPs gleaned from the Illumina Race/Ethnicity panel (see below). We have mapped these sequences to the mitochondrial genome (see figure 3 below) to demonstrate the distributions of the mutations found in the various mitochondrial genes. Mutations below the 0.01 variant ratio have been filtered out. Figure 3A shows the SNVs in the mitochondrial genes. As expected the most mutations are seen in the D-loop, but a significant number of additional mutations are found in other regions of the genome as well including the coding regions of for a number of components of the electron transport chain. Similar data in Figure 3B shows the distribution of insertions and deletions across the mitochondrial genes.

Figure 3: Distribution of variants found in mitochondrial genes for the initial subjects.

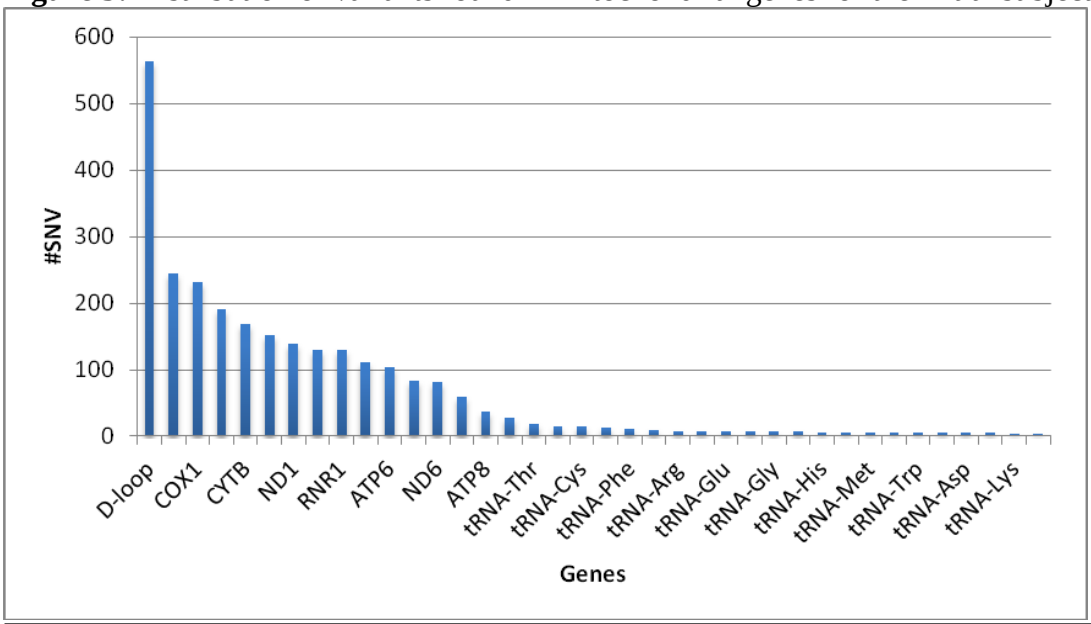


Figure 3A: Distribution of mitochondrial single nucleotide variants found within various mitochondrial genes for the initial subjects as compared to the Cambridge reference sequence. The Y axis represents the absolute number of variants observed within each gene.

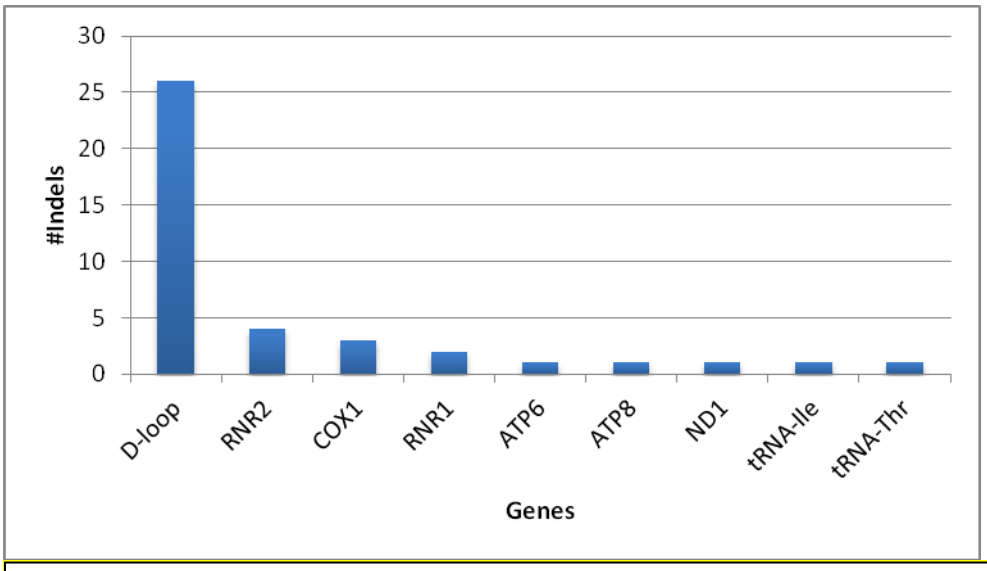


Figure 3B Unique insertions and deletions by gene across the initial patients as compared to the Cambridge reference sequence. The Y axis represents the absolute number of insertions/deletions observed within each gene.

Although not covered by this project, we have also sent more than 300 samples for genotyping on the Illumina African American Admixture panel. Thus far all samples have been successfully genotyped.

<u>Aim 2:</u> To understand the role of mitochondrial DNA variants in regulation of cellular processes important in the cancer cell including those involved with generation of reactive oxygen species and energy metabolism.

Task 1: <u>Creation of Cybrid cell lines</u>

1a.	*IRB submission for creation of lymphoblastoid lines from patients*	*Months 1-4*
1b.	*Creation of immortalized lymphocyte bank for use in study*	*Months 5-24*
1c.	*Order/test primers for variable number tandem repeat studies of nuclear origin*	*Months 6-12*
1d.	*Deplete prostate cancer cells of mitochondria 3 potential cell lines to be used: PNT1A, LNCaP or PC3*	*Months 9-24*
1e.	*Create cytoplasts and fuse with prostate cancer cells devoid of mitochondria to create cybrid cell lines*	*Months 20-36*
1f.	*Confirm nuclear origin of cybrid cell lines using variable length tandem repeat in insulin receptor and other genes*	*Months 20-36*

Task 2: <u>Testing of cybrid cells lines</u>

2a.	*Proliferation and cell cycle testing, Oxygen consumption, Production of reactive oxygen species, electron transport chain analysis, analysis of anchorage-independent cell growth.*	*Months 22-40*

Initially we encountered a number of technical difficulties during the generation of cybrid models using the ethidium bromide (EBr)-mediated mitochondrial DNA (mtDNA) depletion method. During the past year, we continued our original cybrid protocol using the EBr method by varying treatment conditions. Unfortunately, despite our best efforts, quantitative real time PCR (qPCR) analysis suggested a lack of complete mtDNA depletion in these cell lines (rho zero cells). Without the rho zero cells, we could not generate cybrids in prostrate cancer models using the original protocol. To overcome these difficulties, we tested multiple alternative approaches including rhodamine-6G (R6G) mediated short-term mitochondrial dysfunction in generating rho zero cells. We used the benign prostrate cell line PNT1A and the prostate cancer cell line PC3. Generation of rho zero using another prostate cancer cell line, LNCaP, was not successful due to low adherence of these cells to the bottom of cell culture plates after treatment with mtDNA depletion agents. Lack of adherence has made separation of viable rho zero cells very difficult.

Repeated efforts using R6G treatment finally succeeded in the generation of PNT1A rho zero; we have not yet been successful with the PC3 cells, which did not survive the rho zero condition. Cybrid models were then generated in the PNT1A benign prostate nuclear background. Mitochondria from enucleated metastatic prostate cancer cell line PC3 or benign prostrate cell line PNT1A were fused with the rho zero PNT1A nuclear donor to form the cybrids (PC3/PNT1A and PNT1A/PNT1A respectively). We did extensive analysis of these cybrids to confirm their mitochondrial and nuclear identity. We sequenced multiple regions of mtDNA to confirm the mitochondrial identity (Figure 1). Similarly, we sequenced several nuclear SNP markers (Figure 2) as well as short tandem repeat **(**STR**)** markers (Figure 3) to confirm the nuclear identity of the cybrids. These analyses confirmed that both PC3/PNT1A and PNT1A/PNT1A cybrids contain PNT1A benign nuclear background but mitochondria from PC3 and PNT1A, respectively.

In order to confirm the role of mitochondrial retrograde regulation of cancer mitochondria in a nuclear background we analyzed the cancer property of these cybrids. We used the soft agar colony formation assay with similar number of PC3 and PNT1A parental cell lines as well as PC3/PNT1A and PNT1A/PNT1A cybrids under uniform experimental conditions. We observed that the while PNT1A/PNT1A cybrid which contain both nucleus and mitochondria from the benign cell line rarely form colonies in the soft agar like PNT1A parental

cells, PC3/PNT1A cybrid with cancer mitochondria could grow anchored dependently in the soft agar (Figure 4).

Figure-1: Confirmation of mitochondria by sequencing mtDNA of parental cells and cybrids

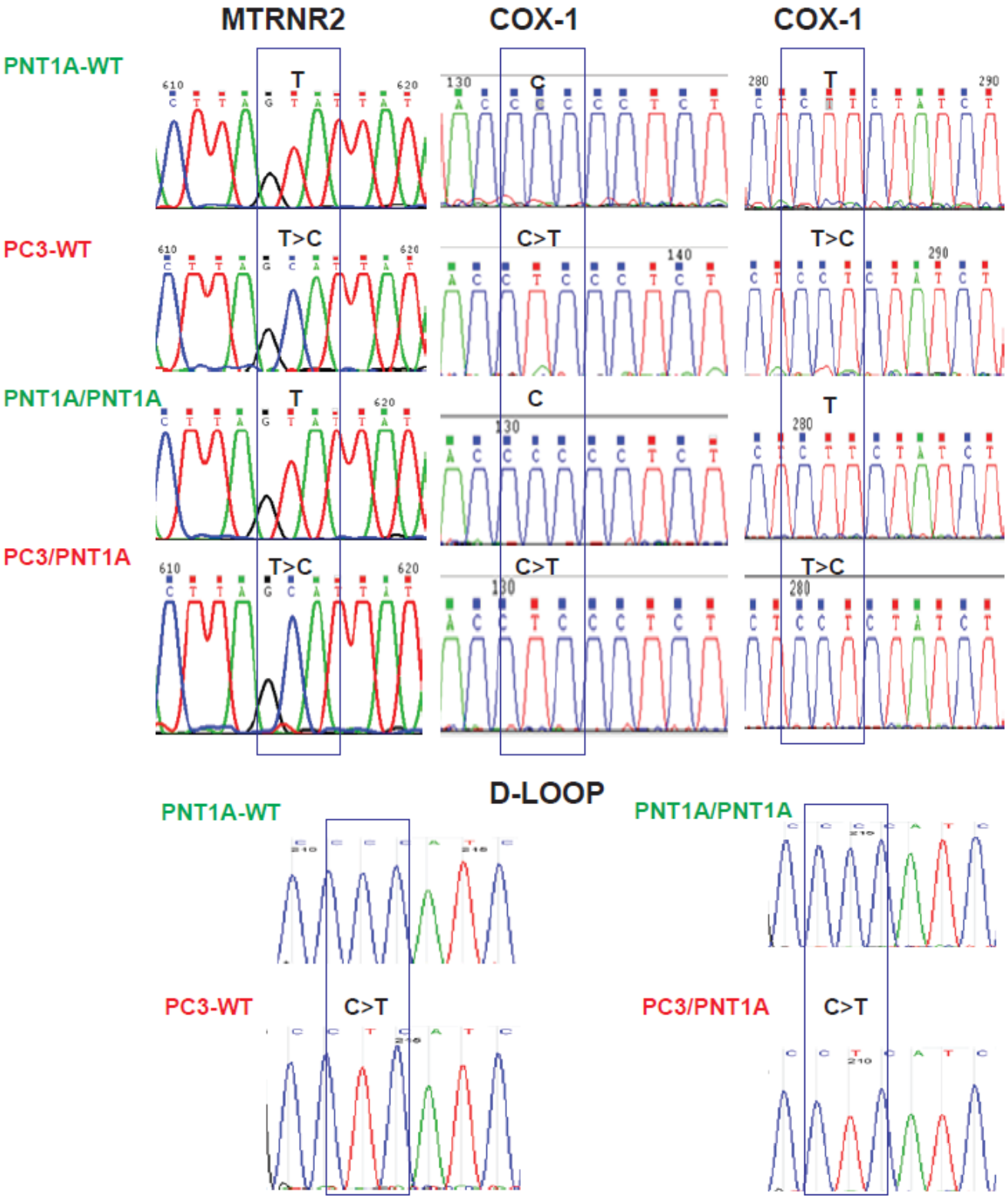


Figure-2: Confirmation of nucleus by sequencing random SNPs in nuclear DNA

PNT1A-WT

No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PNT1A-InQC-1F_2013-10-22_B01.ab1	1-F	InternalQC-01	chr1	-55	58	111	14	1	c.2C>GA,p.S1WX
1							111	14	1	100.00%
2	PNT1A-InQC-3F_2013-10-22_B02.ab1	2-F	InternalQC-03	chr8	-40	58	98	17	1	c.2G>CT,p.R1TM
1							98	17	1	100.00%
3	PNT1A-InQC-5F_2013-10-22_B03.ab1	3-F	InternalQC-05	chr11	-55	49	104	20	1	c.2G>CT,p.R1TM
1							104	20	1	100.00%
4	PNT1A-InQC-7F_2013-10-22_B04.ab1	4-F	InternalQC-07	chr13	-55	58	113	31	1	c.2G>CT,p.R1PL
1							113	31	1	100.00%
5	PNT1A-InQC-9F_2013-10-22_B05.ab1	5-F	InternalQC-09	chr15	-54	51	105	9	1	c.2T>A,p.I1N
1							105	9	1	100.00%
6	PNT1A-InQC-10F_2013-10-22_B06.ab1	6-F	InternalQC-10	chr19	-10	58	68	10	1	c.2G>CT,p.R1TM
1							68	10	1	100.00%
7	PNT1A-InQC-12F_2013-10-22_B07.ab1	7-F	InternalQC-12	chr20	-51	55	106	35	1	c.2T>GA,p.L1RQ
1							106	35	1	100.00%

PC3-WT

No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PC3-InQC-1F_2013-10-22_C01.ab1	1-F	InternalQC-01	chr1	-55	58	111	20	1	c.2C>GA,p.S1WX
1							111	20	1	100.00%
2	PC3-InQC-3F_2013-10-22_C02.ab1	2-F	InternalQC-03	chr8	-40	58	98	22	1	c.2G>TC,p.R1MT
1							98	22	1	100.00%
3	PC3-InQC-5F_2013-10-22_C03.ab1	3-F	InternalQC-05	chr11	-55	49	104	18	1	c.2G>CT,p.R1TM
1							104	18	1	100.00%
4	PC3-InQC-7F_2013-10-22_C04.ab1	4-F	InternalQC-07	chr13	-55	58	113	29	1	c.2G>CT,p.R1PL
1							113	29	1	100.00%
5	PC3-InQC-10F_2013-10-22_C08.ab1	5-F	InternalQC-10	chr19	-10	57	87	18	1	c.2G>T,p.R1M
1							87	18	1	100.00%
6	PC3-InQC-12F_2013-10-22_C07.ab1	6-F	InternalQC-12	chr20	-52	55	107	31	1	c.2T>GA,p.L1RQ
1							107	31	1	100.00%
7	PC3-InQC-9F_2013-10-22_C05.ab1				-1	-178	179	0	-1	Bad Data

PNT1A/PNT1A

No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PNT1A-InQC-1F_2013-10-22_B01.ab1	1-F	InternalQC-01	chr1	-55	56	111	14	1	c.2C>GA,p.S1WX
1							111	14	1	100.00%
2	PNT1A-InQC-3F_2013-10-22_B02.ab1	2-F	InternalQC-03	chr6	-40	58	98	17	1	c.2G>CT,p.R1TM
1							98	17	1	100.00%
3	PNT1A-InQC-5F_2013-10-22_B03.ab1	3-F	InternalQC-05	chr11	-55	49	104	20	1	c.2G>CT,p.R1TM
1							104	20	1	100.00%
4	PNT1A-InQC-7F_2013-10-22_B04.ab1	4-F	InternalQC-07	chr13	-55	58	113	31	1	c.2G>CT,p.R1PL
1							113	31	1	100.00%
5	PNT1A-InQC-9F_2013-10-22_B05.ab1	5-F	InternalQC-09	chr15	-54	51	105	9	1	c.2T>A,p.I1N
1							105	9	1	100.00%
6	PNT1A-InQC-10F_2013-10-22_B06.ab1	6-F	InternalQC-10	chr19	-10	58	68	10	1	c.2G>CT,p.R1TM
1							68	10	1	100.00%
7	PNT1A-InQC-12F_2013-10-22_B07.ab1	7-F	InternalQC-12	chr20	-51	55	106	35	1	c.2T>GA,p.L1RQ
1							106	35	1	100.00%

PC3/PNT1A

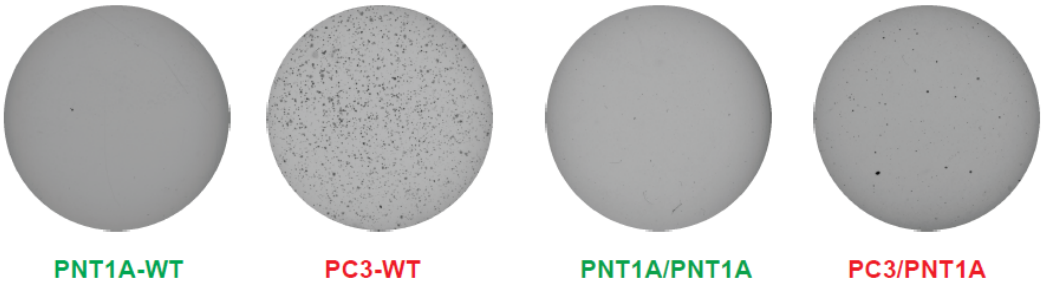
No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PC_PNT1A-PLK01-InQC-1F_2013-10-22_E01.ab1	1-F	InternalQC-01	chr1	-55	56	111	28	1	c.2C>GA,p.S1WX
1							111	28	1	100.00%
2	PC_PNT1A-PLK01-InQC-3F_2013-10-22_E02.ab1	2-F	InternalQC-03	chr8	-40	58	98	23	1	c.2G>TC,p.R1MT
1							98	23	1	100.00%
3	PC_PNT1A-PLK01-InQC-5F_2013-10-22_E03.ab1	3-F	InternalQC-05	chr11	-55	49	104	21	1	c.2G>CT,p.R1TM
1							104	21	1	100.00%
4	PC_PNT1A-PLK01-InQC-7F_2013-10-22_E04.ab1	4-F	InternalQC-07	chr13	-52	58	110	28	1	c.2G>CT,p.R1PL
1							110	28	1	100.00%
5	PC_PNT1A-PLK01-InQC-9F_2013-10-22_E05.ab1	5-F	InternalQC-09	chr15	-50	51	101	22	1	c.2T>A,p.I1N
1							101	22	1	100.00%
6	PC_PNT1A-PLK01-InQC-10F_2013-10-22_E06.ab1	6-F	InternalQC-10	chr19	-13	57	70	18	1	c.2G>CT,p.R1TM
1							70	18	1	100.00%
7	PC_PNT1A-PLK01-InQC-12F_2013-10-22_E07.ab1	7-F	InternalQC-12	chr20	-50	55	105	38	1	c.2T>GA,p.L1RQ
1							105	38	1	100.00%

Figure 3 STR profiles of cybrids and parental cells

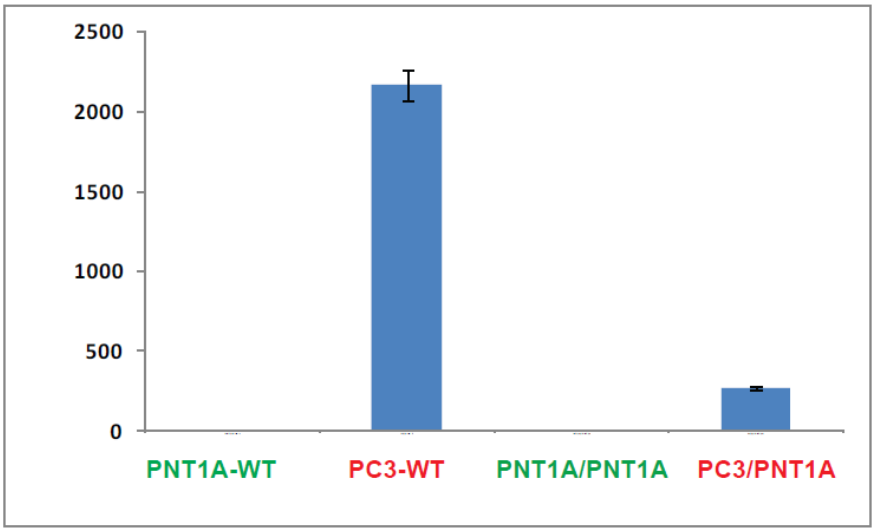
Sample_Name	AMEL	CSF1PO	D13S317	D16S539	D18S51	D21S11	D3S1358	D5S818	D7S820	D8S1179	FGA	TH01	TPOX	vWA
PNT1A-WT	X,Y	12	11,12	10,11	16	27,28	16	12	10,11	10,12	23,24	6,7,9.3	8,11	17,19
PC3/PNT1A-PLKO1	X,Y	12	11	10,11	16	27,28	16	12	10,11	10,12	24	6,7,9.3	8	17,19
PNT1A/PNT1A-PLKO1	X,Y	12	11,12	10,11	16	27	16	12	10,11	10,12	24	6,7,9.3	8,11	17,19
PC3-ATCC	X	11	11	11				13	8,11			6, 7	8,9	17

Figure-4: Analysis of tumor properties in parentals and cybrids by colony formation assay

4A: Phase contrast images of soft agar colonies after two weeks of culture



4B: Number of colonies counted using Gelcount colony counter



Overall, we successfully generated cybrid models under benign prostate nuclear background and analyzed its genetic and tumor properties using different mitochondrial donors. We are continuing our efforts to generate new cybrid models with cancer nuclear background and to further characterize the already generated PNT1A cybrids.

References:

1. Chen TJ, Boles R, Wong, L-J C. Detection of mitochondrial DNA mutations by temporal temperature gradient gel electrophoresis. *Clin Chem*. 1999; *45:* 1162–1167.
2 Wong LJ, Liang MH, Kwon H, Park J, Bai RK, Tan DJ. Comprehensive scanning of the entire mitochondrial genome for mutations. *Clin Chem.* 2002; 48:1901-12.
3. Zhang W, Cui H, Wong L-J. Comprehensive One-Step Molecular Analyses of Mitochondrial Genome by Massively Parallel Sequencing. Clin Chem. 2012; 58: 1322-1331.
4. Falk MJ, Pierce EA, Consugar M, Xie MH, Guadalupe M, Hardy O, Rappaport EF, Wallace DC, LeProust E, Gai X. Mitochondrial disease genetic diagnostics: optimized whole-exome analysis for all MitoCarta nuclear genes and the mitochondrial genome. Discov Med. 2012 14:389-99.
5. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Sequence and organization of the human mitochondrial genome. Nature 1981; 290: 457–465.
6. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N "Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. Nat Genet 1999; 23: 147.

Key Accomplishments Summarized:

1. Isolation of more than 1500 DNA samples required for the project.
2. Validation of a more robust sequencing technique using single amplicon long-range PCR that permits deep coverage (10,000-20,000X on average) of the mitochondrial genome, to date 652 samples of Frozen Tissue DNA have been sequenced.
3. Sequencing of FFPE DNA samples using Sequence Capture.
4. Mapping of the DNA variants to mitochondrial genes (data shown for ~first 200 samples nearly completed for remaining samples).
5. Cybrid models of benign (PNT1A) prostate cells have been successfully created and characterized. Despite the end of the project funding, we are continuing to work with the PC3 cells with the expectation that we will be able to create PC3 cybrids and characterize them.

Opportunities for professional development:

Though not funded by the DOD directly, several students and postdoctoral fellows have interacted with the project and in the summer of 2015, an undergraduate student from an underserved minority group. This last summer the undergraduate student (under the supervision of one of a postdoctoral fellow in Dr. Wheeler's group) worked on coverage analysis of mitochondrial samples comparing the coverages obtained from FFPE samples to those of fresh frozen samples. He determined that the overall coverage patterns of fresh frozen were different from those of FFPE and that this would require separate analysis of the two approaches. He also identified a region of the FFPE capture that was poorly covered compared with fresh frozen and attempted to map the mitochondrial genes to that area. This area is largely "uncharted territory" and allowed both the postdoctoral fellow and the undergraduate to develop new skills in analyzing sequence data. Post doctoral fellows who were part of the Houston REACH IRACDA, or Houston Research Education and Career Horizon Institutional Research and Academic Career Development Award, K12 Postdoctoral Fellowship Program have also participated in the project. This IRACDA program is a model for postdoctoral training that combines frontier-level biomedical research and training to teach research-oriented courses to undergraduates at partner institutions. The project includes under-represented post-doctoral fellows when possible, who provide excellent role models for the undergraduates.

Dissemination of data to communities of interest:
Though data from this work has not yet been published in a peer reviewed journal, portions of the data have been presented in local venues including the Human Genetics Postdoctoral seminar series and the SMART (Summer Medical And Research Training) Program Seminar series aimed at undergraduate students interested in research and medical careers.

Next Reporting period plans:
Nothing to report (final report)

4. Impact:

Development of the principal discipline:
This project performed mitochondrial sequencing of more than 1000 subject samples from African American men. To our knowledge this is the largest such database in existence. In concert with Dr. David Wheeler in the Human Genome Sequencing Center and Dr. Lee-Jun Wong, we have implemented a new more rapid and much more powerful technique for sequencing mitochondrial DNA using single amplicon long-range PCR. This technology permitted us to rapidly sequence 652 samples using NGS technology at a depth of coverage of 10,000-20,000X. This technology allows much more accurate base calls and improves the limit of detection for heteroplasmy from about 15% with Sanger sequencing to ~ 1-2% with the long range primer technique. This technology has permitted much more rapid generation of mtDNA sequence with much less DNA and less "contamination" by nuclear DNA. Further, we have optimized sequencing strategies for the FFPE DNA and this portion of the project has just been completed. Two strategies were necessary for sequencing these two kinds of DNA samples and both of these strategies were either developed or perfected during the course of this project. Publication of the methodology and the data obtained will represent a huge leap forward for the field of mitochondria and cancer. Further, we have created a large database of DNA sequence combined with clinical data and we are in the process of analyzing this data to produce the first manuscript relating to this work. We envision that this annotated data will reveal that mitochondrial metabolism directly impacts the development of prostate cancer.

Impact on other disciplines:
Though we had not foreseen this when the project was first envisioned, we are now working with investigators in metabolomics to examine whether the mitochondrial sequence changes we identify can be correlated with changes in metabolism in vivo (urine or plasma metabolome).

Impact on Technology Transfer:
Nothing to report

Impact on Society:
The overarching goal of this project is to understand why African American men are more prone to develop prostate cancer than any other racial/ethnic group. We envision that mitochondrial inheritance plays a significant role in aggressiveness of prostate cancer in African American men. Uncovering risk alleles may impact both screening and treatment for these men.

5. Changes/Problems:

Changes in Approach:
No significant changes in objectives or scope were encountered. As detailed in Aim 1, improvements in technology developed during the project allowed us to gain much deeper sequencing information than originally envisioned. As detailed in Aim 2, the first approach to developing cybrids was not feasible and we utlilized an alternative technology which was successful in the PNT1A cells.

Problems or Delays
Creation of the cybrid cells proved to be more difficult than we had envisioned at the outset of the project. Ultimately we were able to develop and characterize the PNT1A cybrids, but we are still struggling the PC3

cybrids making slow progress. As noted below, we will be able to share our insights in a methods paper which we hope to publish in the spring of 2016.

Changes that had a significant impact on expenditures
None

Significant changes in humans subjects, vertebrate animals or select agents:
None

6. Products:
Publications:
No completed reportable outcomes have been accomplished as yet, however; the mitochondrial database with mtDNA sequence data, African American Admixture Data and Clinical Data is being assembled now. This is a massive amount of data – more than we had originally foreseen. In order to analyze the data, we have recruited Dr. Aaron Thrift, in the Cancer Epidemiology program at the Dan L Duncan Cancer Center. With the help of Dr. Thrift, we will be in a position to submit the first manuscript on this data this spring of 2016 – later than we had wanted to do, but we have overcome the technical challenges inherent in sequencing the FFPE DNA samples to create a very large database of information. The cybrid cell lines were unfortunately more difficult to create than we had anticipated, but we have succeeded in creating PNT1A cybrids and we have not given up on the PC3 cybrids.

Websites:
None

Technologies:
Technology for sequencing mitochondrial DNA from FFPE samples is described above. In concert with Dr. David Wheeler in the Human Genome Sequencing Center and Dr. Lee-Jun Wong, we have implemented a new more rapid and much more powerful technique for sequencing mitochondrial DNA using single amplicon long-range PCR. This technology permitted us to rapidly sequence 652 samples using NGS technology at a depth of coverage of 10,000-20,000X. This technology allows much more accurate base calls and improves the limit of detection for heteroplasmy from about 15% with Sanger sequencing to ~ 1-2% with the long range primer technique. This technology has permitted much more rapid generation of mtDNA sequence with much less DNA and less "contamination" by nuclear DNA. We have optimized sequencing strategies for the FFPE DNA and this portion of the project has just been completed. We will publish this information in a peer reviewed journal along with the actual data gleaned from these samples. The pitfalls of using DNA derived from FFPE samples as compared to DNA derived from frozen tissue will be discussed in detail as will details of how to manage these kinds of data.

Inventions:
None

Other Products:
We are in the process of creating a very large database of DNA sequence information and clinical data on African American men with prostate cancer. This information will be invaluable in terms of evaluating the role mitochondrial inheritance on prostate cancer aggressiveness. Further, we have developed a library of germline DNA which can be queried for other projects.

7. Participants

Martha Mims, MD, PhD, PI No Change
Lee-Jun Wong, PhD, Co-Investigator No Change

David Wheeler, PhD, Co-Investigator No Change
Suzanne Leal, PhD, Co-Investigator No Change
Myoung Kwon, Research Technician, No Change

8. Special Reporting Requirements: None

9. Appendices: None